AFHRL-TP-86-37

12

AIR FORCE HUMAN RESOURCES LABORATORY

JOB PERFORMANCE MEASUREMENT: A SYSTEMATIC PROGRAM OF RESEARCH AND DEVELOPMENT

Jerry W. Hedge
Mark S. Teachout

TRAINING SYSTEMS DIVISION
Brooks Air Force Base, Texas 78235-5601

November 1986
Interim Technical Paper for Period June 1985 - April 1986

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

AD-A174 175

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFHRL-TP-86-37	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Training Systems Division 6b. OFFICE SYMBOL (*If applicable*) AFHRL/IDE	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (*City, State, and ZIP Code*) Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235-5601	7b. ADDRESS (*City, State, and ZIP Code*)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Air Force Human Resources Laboratory 8b. OFFICE SYMBOL (*If applicable*) HQ AFHRL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (*City, State, and ZIP Code*) Brooks Air Force Base, Texas 78235-5601	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 62703F; PROJECT NO. 7734; TASK NO. 08; WORK UNIT ACCESSION NO. 22

11. TITLE (*Include Security Classification*)
Job Performance Measurement: A Systematic Program of Research and Development

12. PERSONAL AUTHOR(S)
Hedge, J.W.; Teachout, M.S.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (*Year, Month, Day*)	15. PAGE COUNT
Interim	FROM Jun 85 TO Apr 86	November 1986	20

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (*Continue on reverse if necessary and identify by block number*)
FIELD	GROUP	SUB-GROUP	criterion development job performance performance measurement
05	09		
05	10		

19. ABSTRACT (*Continue on reverse if necessary and identify by block number*)

This paper describes the Air Force's strategy for developing, evaluating, and selecting a measurement methodology for use in validating selection systems and evaluating training programs across a wide range of job specialties. Background, current status, and future plans of the job performance measurement project are discussed in terms of alleviating the criterion problem that has defied solution by industrial psychologists for many years.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Nancy A. Perrigo, Chief, STINFO Office	22b. TELEPHONE (*Include Area Code*) (512) 536-3877 22c. OFFICE SYMBOL AFHRL/TSR

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

AFHRL Technical Paper 86-37

November 1986

JOB PERFORMANCE MEASUREMENT: A SYSTEMATIC PROGRAM OF RESEARCH AND DEVELOPMENT

Jerry W. Hedge
Mark S. Teachout

TRAINING SYSTEMS DIVISION
Brooks Air Force Base, Texas 78235-5601

Reviewed and submitted for publication by

Roger D. Ballentine, Lt Colonel, USAF
Chief, Skills and Development Branch
Training Systems Division

This publication is primarily a working paper. It is published solely to document work performed.

SUMMARY

This paper describes the Air Force's ongoing job performance criterion development project. Specifically, details are presented on the background and conceptual basis for this work, as well as the job performance measurement methodology that has been developed to date. In addition, the strategy being used to identify a measurement system across clusters of specialties is discussed. Finally, ongoing and planned research and development efforts which support this project are described.

PREFACE

Approximately 4 years have passed since the initiation of a long-term research and development program focusing on job performance criterion development. This paper is intended as an interim description of the project - its past, present, and future. An earlier version of this paper was presented at the First Annual Conference of the Society for Industrial and Organizational Psychology, Chicago, Illinois, April 1986.

TABLE OF CONTENTS

		Page
I.	INTRODUCTION	1
II.	CRITERION DEVELOPMENT	2
III.	AIR FORCE JOB PERFORMANCE MEASUREMENT SYSTEM	3
	Walk-Through Performance Testing	3
	Rating Forms	4
IV.	JPMS DEVELOPMENT STRATEGY	5
V.	APPLICATION OF JPMS TECHNOLOGY TO TRAINING AND OFFICER/CIVILIAN SELECTION	7
	Training System Evaluation	7
	Officer/Civilian Selection System Validation	9
VI.	INTEGRATION OF PRIMARY AND SUPPORTIVE R&D	9
VII.	SUMMARY	12
REFERENCES		12

LIST OF FIGURES

Figure		Page
1	JPMS Development Strategy	6
2	Application of JPMS Technology to Training System Evaluation and to Officer/Civilian Selection	8
3	Integration of Primary and Supportive R&D	10

LIST OF TABLES

Table		Page
1	Air Force Specialties Selected for Joint-Service Effort	2
2	Summary of R&D Efforts	11

JOB PERFORMANCE MEASUREMENT: A SYSTEMATIC PROGRAM OF RESEARCH AND DEVELOPMENT

I. INTRODUCTION

The Air Force Human Resources Laboratory (AFHRL) is currently conducting a large-scale effort to develop a measurement technology for systematically obtaining job performance data. Planning for the Air Force's program of research and development (R&D) began several years ago as the result of three primary requirements. As an initial requirement, operational military and civilian program managers in the manpower, personnel, and training communities asked AFHRL to develop an approach for measuring job performance so that measures could be used to assist in the evaluation of their training and selection programs. In addition, the appearance in 1978 of the Uniform Guidelines for Employee Selection highlighted the increasing need to validate civilian selection procedures against job performance measures (Cascio & Bernardin, 1981).

Secondly, the manpower, personnel, and training research community needed performance measures to serve as criteria in their R&D projects. This need was endorsed when an AFHRL Research Advisory Panel (composed of knowledgeable scientists from academia and industry, as well as peers from the Army and Navy) recommended that research efforts to obtain specific performance criteria be consolidated to develop a job performance measurement technology composed of a wide range of candidate measurement systems to serve the varied projects' needs.

Plans for the Air Force job performance measurement effort to meet these requirements were already under development when a third requirement for these measures came with the Congressional mandate to test the feasibility of validating the Armed Services Vocational Aptitude Battery (ASVAB) against job performance measures. The impetus for this mandate occurred in 1980, when Congress was informed that there was a problem with the norming of the ASVAB, and that the Services might be enlisting a greater percentage of low-scoring individuals than was previously thought to be the case. (Norming is the method through which test raw scores are converted into percentile scores that permit them to be compared to the scores of a defined and relevant population. These norms allow the DoD to evaluate its new recruits across time and across the Armed Services.)

As a result of the norming issue, the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics) directed the Military Services to establish an R&D program to link enlistment standards to job performance. In 1983, Congress tasked the Office of Assistant Secretary of Defense to provide direct oversight for joint-Service R&D activities addressing job performance measurement. In addition, DoD requested that the National Academy of Sciences impanel a Committee on the Performance of Military Personnel (composed of nationally recognized experts in scientific and technical areas related to performance measurement). This committee provides an independent, technical review of the joint-Service project. Thus, the operational requirements and the legal and Congressional mandates provided the impetus to planning and obtaining support for a lengthy, high-resource R&D effort.

The Air Force research plan has been developed in coordination with the other Services to ensure that all programs are complementary. An integrated set of cross-Service and Service-specific enlisted specialties has been designated to serve as the initial test bed for the development of job performance measurement strategies. Eight Air Force specialties (AFSs) were selected for the Air Force component of the joint-Service project. These specialties and their aptitude index areas are listed in Table 1.

These AFSs cover the range of jobs in the mechanical, administrative, general, and electronics aptitude areas. Further developmental efforts are planned for additional enlisted

Table 1. Air Force Specialties Selected for Joint-Service Effort

Air Force specialty	Specialty code	Aptitude index area
Jet Engine Mechanic	426X2	Mechanical
Ground Radio Operator	492X1	Administrative
Air Traffic Control Operator	272X0	General
Avionic Communications Specialist	328X0	Electronics
Aerospace Ground Equipment Mechanic	423X5	Mechanical
Personnel Specialist	732X0	Administrative
Aircrew Life Support Specialist	122X0	General
Precision Measuring Equipment Specialist	324X0	Electronics

AFSs from each aptitude area, as well as selected officer and civilian specialties. In addition, because each Service has chosen a unique performance measurement orientation, technologies developed by the Air Force in cross-Service specialties can be transferred to other Services.

The overall program of research calls for the development of measurement techniques that will allow for the collection of valid, accurate, and reliable hands-on job performance information. These measures in turn will be used as benchmarks against which surrogate indices of performance (less expensive, easier-to-administer interview tests, and performance ratings) will be evaluated as substitutes for the more expensive, labor-intensive, hands-on performance measures. In addition to newly developed performance measures, existing measures such as technical training scores, Airman Performance Report ratings, and skill-level advancement indices will also be evaluated as possible surrogate measures.

The rationale for this multi-faceted approach to performance measurement involves cost considerations (the hands-on measures are very expensive to develop and administer), but it is also based on the hypothesis that the various methods - while to some extent overlapping - measure different aspects of job performance with differing levels of accuracy. (See Kavanagh, Borman, Hedge, and Gould (1982) for a detailed discussion of this approach.) The Air Force strategy for this multi-faceted approach will be discussed in more detail in a later section of this paper.

II. CRITERION DEVELOPMENT

The overall purpose of the R&D program discussed in this paper is to develop and evaluate a strategy for using measures of job performance for validation of selection and classification procedures and for evaluating training programs. In order to succeed in this criterion development effort, the job performance measures must constitute a valid, accurate reflection of an individual's work performance.

A variety of alternative measures are available for assessing performance on the job. They range from subjective to objective and from general to specific. When faced with the choice of which criterion to select, the researcher or practitioner typically relies on several informal decision rules:

1. cost in terms of time, money, safety, or mission effects.
2. convenience in developing or obtaining measures.
3. fidelity or accuracy of replicating behaviors relevant to the job.

Unfortunately, the development of a criterion measure is frequently seen as a secondary concern, and Decision Rule 2 - convenience - is frequently applied. The result is a generic, packaged-to-please rating form that will in all probability satisfy Decision Rule 1 but not Rule 3.

When the chief concern of researchers and practitioners shifts to Decision Rule 3, and fidelity becomes the overriding concern, measuring a sample of actual work presents a viable alternative to the convenient and more subjective rating form. The technique of work sample testing involves an individual's performing a task or selected set of tasks relevant to his/her job and the objective scoring of performance on these tasks. The value of the work sample methodology lies in the fidelity with which the selected set of tasks allows measurement of an incumbent's job proficiency. Unfortunately, this can also be a weakness of the technique. Work sample procedures normally identify critical tasks, discard those not practically measurable, and then simply allow the remainder to become the selected set of tasks to be measured. AFHRL's approach to work sample testing is an attempt to overcome this criterion deficiency problem.

III. AIR FORCE JOB PERFORMANCE MEASUREMENT SYSTEM

AFHRL's approach to this criterion problem calls for the development of a variety of measurement techniques. The most detailed method, known as Walk-Through Performance Testing (WTPT), combines the observation of hands-on performance and incumbent interview testing for a set of representative tasks. In addition, task, dimensional, global, and Air Force-wide rating forms have been developed for use by supervisors, peers, and incumbents. All four sets of rating forms require evaluation of job performance. The WTPT and rating forms in combination constitute the Job Performance Measurement System (JPMS). Ultimately, performance information provided by such a system should prove useful for evaluating how well an individual is prepared to accomplish his/her job after completion of the various stages of training (e.g., technical school, field training detachment, or on-the-job).

Walk-Through Performance Testing

For the Air Force, hands-on testing is a particular problem because of the complexity and expense involved in performing many tasks. For example, many critical tasks cannot be measured by hands-on testing because these tasks tend to take too long to complete, require replacement of expensive parts, or risk possible damage to components. AFHRL has developed a new methodology to deal with these problems. This new approach, WTPT, has as its foundation the work sample philosophy but attempts to expand the measurement of critical tasks to include those tasks not measured by hands-on testing, through the addition of an interview testing component (Hedge, 1984).

The hands-on component of the WTPT resembles a traditional hands-on work sample test designed to measure proficiency on a critical task. For example, one hands-on task may require an incumbent to install a starter on a jet engine. On the first page of the examiner's manual, information is provided to the test administrator concerning: testing time; required tools, technical orders, and job guides; pertinent background information and necessary engine configuration; and test administrator instructions. While the starter is being installed, the test administrator uses a checklist to indicate whether steps are performed correctly or not (e.g., lubricate the spline, index the position of the starter, and install the locking device). Finally, a 5-point rating scale allows the test administrator to record an overall rating of proficiency on that task.

Many tasks are either too time-consuming, too costly, or too dangerous to measure by hands-on testing. Interview testing attempts to expand the content domain by measuring tasks that cannot be measured with the hands-on method. Interview testing requires the incumbent to explain the step-by-step procedures necessary for successful completion of the task. This allows the test administrator to assess an incumbent's proficiency-based strengths and weaknesses related to the performance of that task. For example, an interview item may test an incumbent's ability to determine the source of high oil consumption. Once again, on the first page of the administrator's manual, pertinent information is provided to the test administrator. While the incumbent is explaining how to perform the task, the test administrator uses a checklist to indicate whether the steps necessary for successful performance are correctly described. In addition, a 5-point overall proficiency rating is recorded by the administrator.

The interview testing is conducted at the work site in a "show-and-tell" fashion that allows the incumbent to "visually and verbally" describe how a step is to be accomplished (e.g., "that bolt is to be turned five revolutions," or "that component is to be lubricated prior to being assembled"). Thus, information on additional tasks can be collected along with hands-on information to provide a more thorough coverage of the content domain and a more accurate picture of an individual's proficiency.

Rating Forms

Information learned in the WTPT development process was used to develop rating forms that range in detail from the very micro to the very macro. This range provides a thorough coverage of the performance evaluation continuum. These rating forms have been developed as either supplements to the WTPT process or as substitutes for this work sample approach. Four rating forms - task, dimensional, global, and Air Force-wide - comprise this rating system across three separate rating sources (supervisor, peer, and self).

AFHRL's approach to the development of these rating forms involves several key concepts. One requirement is scale comparability across rating forms. All rating forms are constructed using a 5-point, adjectivally anchored rating scale. In addition, whenever possible, specific behavioral descriptions are included to provide detailed information to assist the raters in making accurate judgments. Finally, a minimal competency cutoff is incorporated within each rating scale, such that distinctions can be made between those airmen who meet or fail to meet a specified level of job proficiency.

Task Rating Form. The task rating form provides the most specific rating data. Supervisors, peers, and incumbents are asked to rate proficiency on all tasks representative of the job content domain. Only the 5-point, adjectivally anchored rating scale is utilized for this rating form. The use of specific behavioral descriptors for each task (to describe the different levels of proficiency) is considered impractical, as each task would require an entire page to complete.

Dimensional Rating Form. The dimensional rating form provides the second most specific rating data. Again, supervisors, peers, and incumbents are asked to rate the technical proficiency of first-term airmen across important areas of the job. Potential dimensions are identified through factor analysis of occupational survey data, and scales are developed using input from Subject-Matter Experts (SMEs). In a series of SME workshops, a final set of dimensions is identified, and specific behavioral descriptions are developed for each of the five scalar values. The specific descriptors are developed using a behavioral summary statement (BSS) approach (Borman, 1979), where valid SME-generated behavioral anchors at each level are combined to form paragraph descriptors of that level.

Global Rating Form. A global rating form is developed to collect supervisor, peer, and incumbent overall ratings of proficiency. Two items are generated to cover the job domain: (a) technical proficiency and (b) interpersonal proficiency. Once again, a series of SME workshops is used to generate specific behavioral descriptors for each scalar point on the rating form.

Air Force-Wide Rating Form. The Air Force-wide rating form is developed to be representative of all specialties in the Air Force. This rating form focuses on important performance factors finalized at an SME workshop. Because of this inter-specialty focus, SMEs were resource managers working in the Air Force Military Personnel Center (AFMPC), Assignments Section. Resource managers have oversight responsibilities for different specialties, and their combined knowledge provided the details for constructing a BSS scale that could be used across all specialties.

Additional Information Collected. Another important variable, and one that represents a joint-Service commitment, is job experience. Because the Air Force is the lead Service for researching the job experience/job performance link, data are being collected on: (a) time in service, (b) time in unit, (c) time on present job, (d) number of times each task has been performed, (e) last time each task has been performed, and (f) task-level experience ratings. In this way, an experience composite can be formed and related to performance on the job.

Over and above this joint-Service commitment, AFHRL is collecting data on factors related to performance (e.g., aptitude, training, situational constraints, motivation) and factors related to the measurement of performance (e.g., user reaction data such as accuracy/acceptability of measurement instruments, clarity of instructions).

IV. JPMS DEVELOPMENT STRATEGY

Having detailed the Job Performance Measurement System (JPMS), it is important to explain AFHRL's strategy for developing and testing the measurement system in a variety of specialties. The primary research focus for this effort is the application of the job performance methodology within enlisted career fields. Given that the R&D effort reported in this paper is driven by a multiple measurement philosophy, with the objective of identifying which measure or combination of measures provides the optimal criterion for validating the ASVAB, a more thorough explication of this strategy is required. In addition, it should be kept in mind that the ultimate objective is identification of surrogate measures that can be substituted for the more expensive, time-consuming, hands-on procedure.

Although use of surrogates is the desired outcome of this R&D strategy, decisions must be based on analysis of the full range of measurement instruments in the JPMS. Figure 1 depicts the JPMS development strategy that guides the job performance measurement project.

The initial phase of this effort involves JPMS development and analysis in eight AFSs. The JPMS methodology was first developed for the Jet Engine Mechanic specialty and has now been applied to three additional specialties. Application to the final four AFSs will occur during the 4th quarter of FY86. An important pattern to note with regard to these first eight AFSs is the choice of specialties from the four aptitude index (AI) areas of the ASVAB (mechanical, administrative, general, and electronics). Development and testing are done in a cyclical fashion such that each AI is covered before the process is repeated. This ensures representative coverage of Air Force jobs. This cyclical pattern will remain constant throughout the testing process.

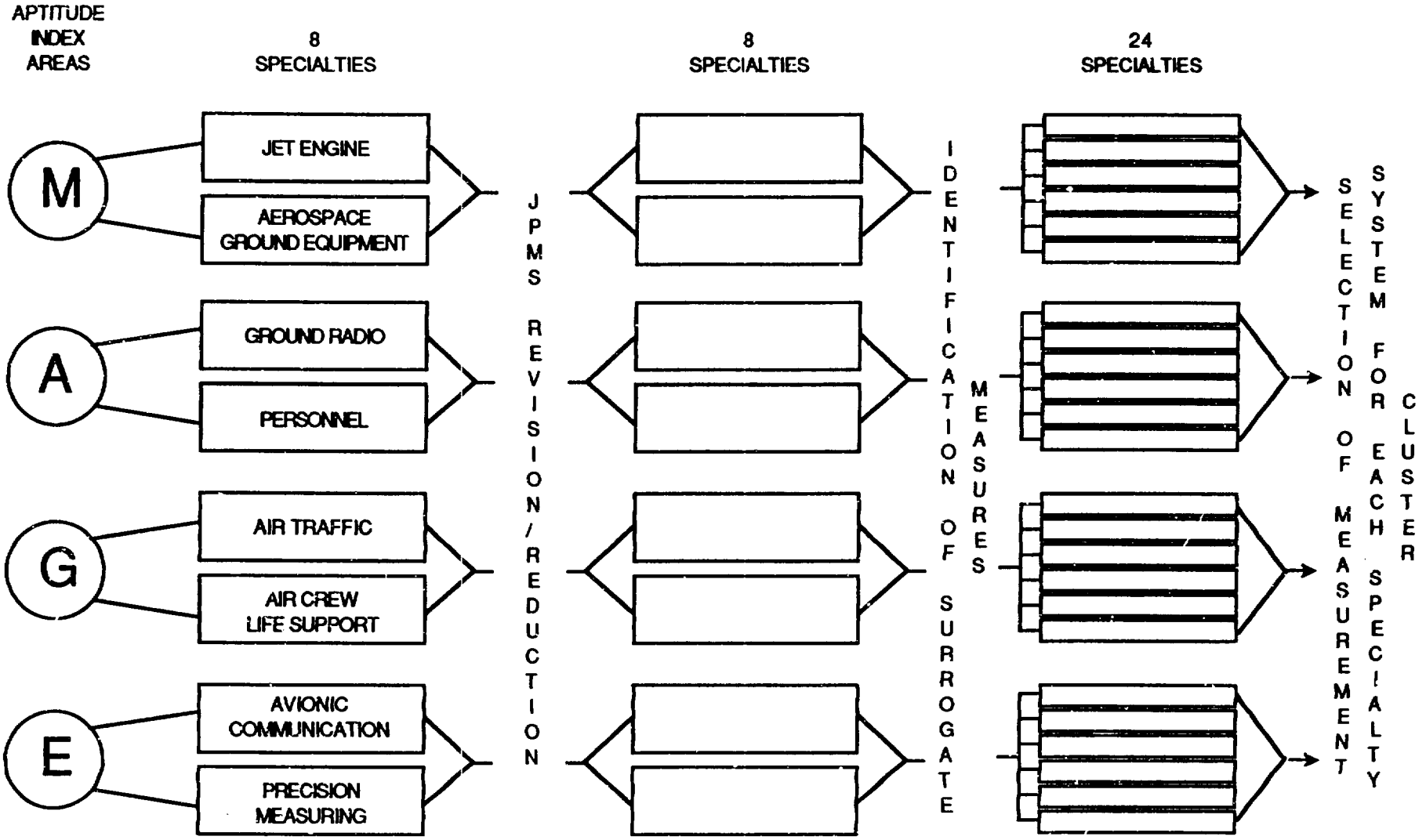

Figure 1. JPMS Development Strategy.

Once development and analysis have been completed on all eight specialties, decisions will be made about revision/reduction of the JPMS prior to application of the system to eight additional specialties. These decisions may involve fine-tuning of JPMS instruments or related materials (e.g., revision of instructions, changes in rater training, changes in rating form layout or format) or reduction in JPMS instruments (e.g., after eight AFSs, it may be determined that self-ratings provide no useful information and should be eliminated). During the second phase, it is still desirable to include hands-on testing as the benchmark for surrogate analysis. In order to identify a meaningful pattern of surrogate instruments and rating sources that adequately cover the criterion space, and to provide accurate measurement of job proficiency, the hands-on benchmark should be included in the JPMS as long as it is feasible. However, due to budgetary constraints, some reduction in the depth and breadth of hands-on testing may be required.

Once the JPMS has been tested in 16 AFSs (four per AI area), cross-specialty results will be analyzed in order to identify patterns of instrument/source superiority. In other words, an attempt will be made to detect clusters of specialties that are defined by surrogate instrument commonality (e.g., supervisor task ratings may be superior surrogates for one cluster, whereas a combination of peer and self dimensional ratings may be more useful for another cluster of specialties). A more concrete example may help to clarify this JPMS development strategy. One possible clustering pattern might be associated with aptitude index areas. That is, perhaps one group of instruments/sources might work best for specialties with a mechanical AI, whereas a second group of instruments/sources might be better suited for administrative specialties. However, rather than hypothesizing clusters and associated instruments, these patterns will be determined empirically.

Finally, based on the analyses for these 16 specialties, appropriate JPMS components will be selected and applied to 24 additional AFSs and used as criteria for validating the ASVAB. Once data have been collected and analyzed for these final 24 AFSs, the JPMS methodology will have been demonstrated in 40 enlisted specialties. Sufficient data should then be available to identify the necessary measurement instruments to validate the enlisted selection test for any specialty.

V. APPLICATION OF JPMS TECHNOLOGY TO TRAINING AND OFFICER/CIVILIAN SELECTION

The previous section of this paper detailed the JPMS development strategy for validation of the ASVAB. AFHRL's plans for application of this methodology to other manpower, personnel, and training requirements will now be discussed.

Training System Evaluation

Initial JPMS work has focused on enlisted selection system validation, as depicted in Figure 2. However, as noted earlier, several other requirements have driven AFHRL's criterion development effort. Primary among these additional foci is training program evaluation. The Air Force conducts extensive training of new enlistees at various phases of their enlistment (employment) (i.e., technical training school, field detachment training, and on-the-job training (OJT)). Preliminary criterion work will concentrate on the potential of the JPMS methodology in evaluating technical school training and OJT.

Most specialties in the Air Force provide initial training once recruits have been classified into a career field. After completing basic military training, recruits are sent to a technical training school for their specialty for a specified period of time (length of training varies by

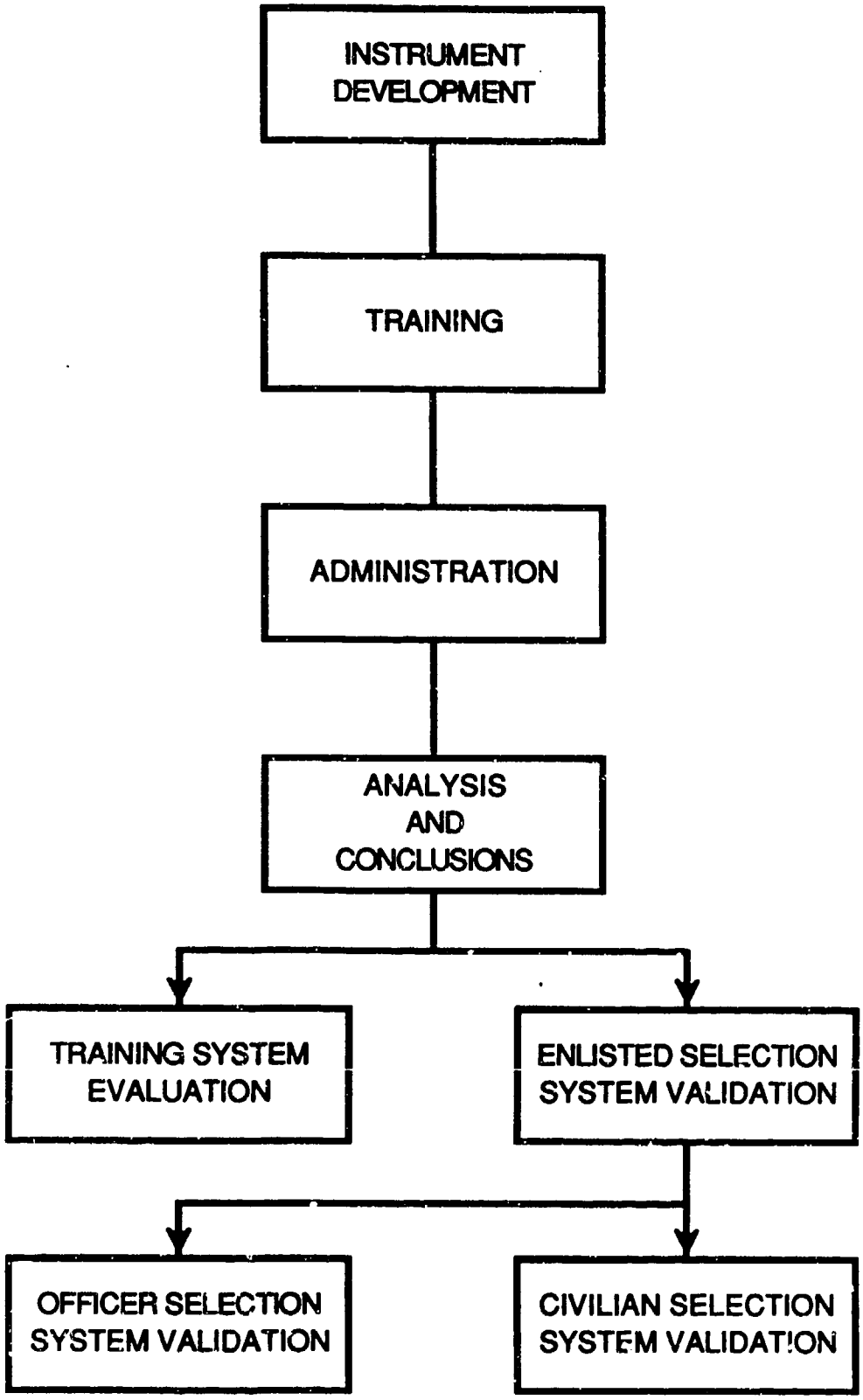

Figure 2. Application of JPMS Technology to Training System Evaluation and to Officer/Civilian Selection.

AFS). Training is given in blocks of instruction, and learning is evaluated with a paper-and-pencil test. Success in technical training is then assessed by averaging scores across blocks of instruction. The JPMS methodology can be utilized to evaluate more accurately performance in technical school.

Once technical school graduates have been placed in their first job assignment, the JPMS methodology can then be applied to assess their proficiency on the job. Once again, individual strengths and weaknesses can be identified, fed back to the individuals, and fed back to the training community. This information provides an opportunity for both technical school and OJT trainers to assess the strengths and weaknesses of their training programs. Work in the training arena is in the planning stages and due to begin in FY87.

Officer/Civilian Selection System Validation

The JPMS methodology will also be transitioned to the officer and civilian selection domains. Just as job performance measures are needed to validate the enlisted selection system, a similar approach may be applied to the officer ranks; however, it is expected that this transition will not occur without some changes to the JPMS. Although the theory and methodology will still be applicable, certain refinements to the JPMS will probably be required. Given the nature of officer work, it is anticipated that the WTPT component of the JPMS may more closely resemble an assessment center design. Development of officer performance measures is planned for FY87.

Similarly, the Air Force's civilian selection system is in need of job performance measures to serve as criteria for validation. A feasibility study to assess the usefulness of the JPMS methodology for measuring the job performance of Air Force civilians is planned for the 4th quarter of FY86. Development and implementation of job performance measures in one civilian job are scheduled for FY87.

Extending the JPMS methodology to officer and civilian jobs is dependent on continued funding at current project funding levels.

VI. INTEGRATION OF PRIMARY AND SUPPORTIVE R&D

In addition to the JPMS development work which has been underway since FY84, a number of ongoing and planned R&D efforts support this primary work effort. Figure 3 describes visually how the results of these efforts are being integrated into the validation project. Rather than describing each research component and how it supports the JPMS, several example studies will be discussed.

One R&D effort currently in progress concerns rater training, as shown in Figure 3. This effort has been designed to address training issues for walk-through test administrators, supervisors, peers, and incumbents who will be completing the four types of rating forms in the JPMS. The interaction of type of training and type of rating form is being investigated, as are the four major components of the behavior modeling paradigm (Sorcher & Goldstein, 1972). The overall test bed for examining these issues employs an assessment center methodology, whereby both observation/recording components (as required for WTPT) and ratings (as required for rating forms in the JPMS) are included. In addition, videotape technology is being used as a mechanism for gathering accuracy data. Finally, the development of an assessment center methodology should prove helpful when developing the WTPT for officer jobs, as well as certain other enlisted specialties (e.g., Personnel).

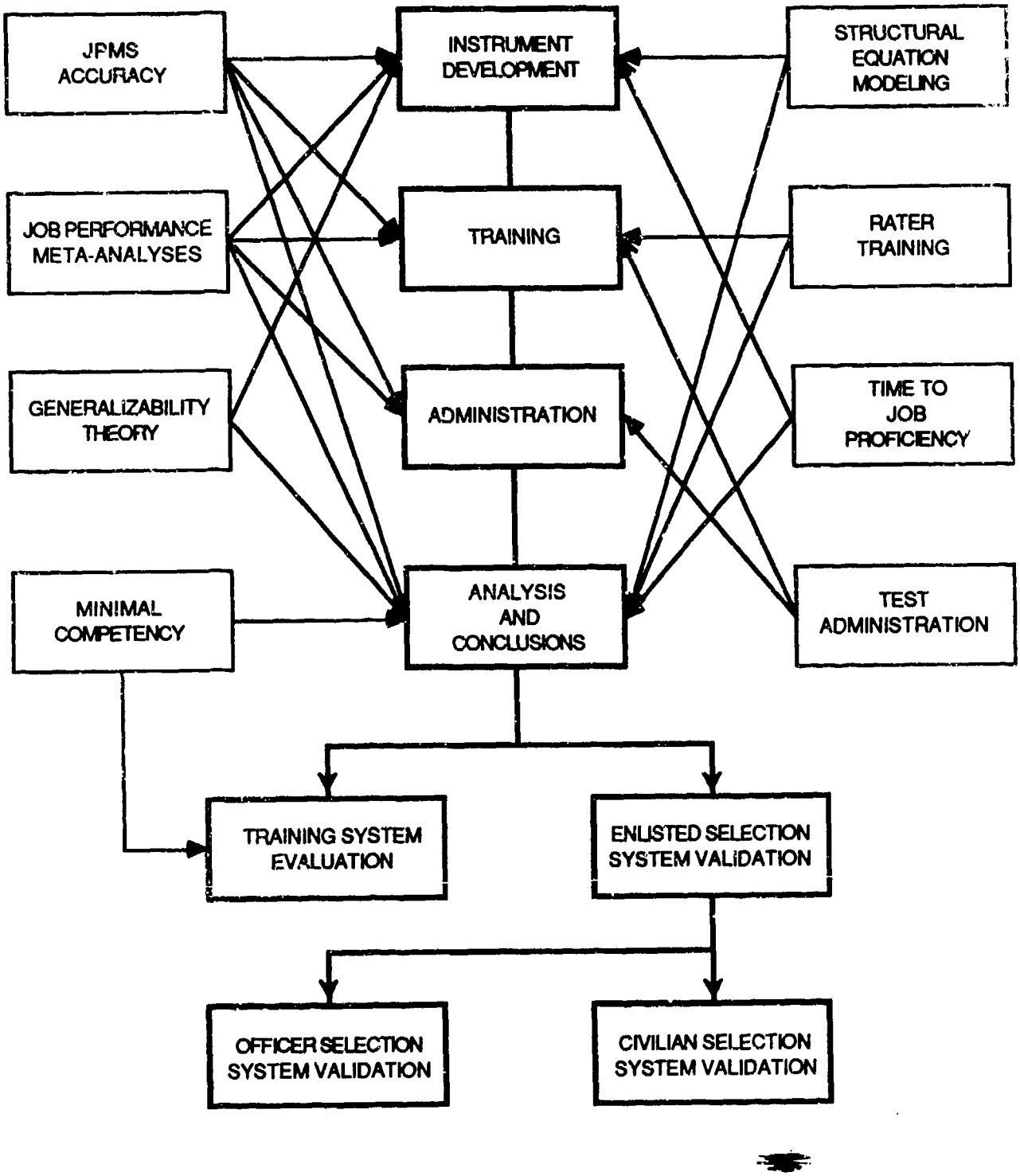

Figure 3. Integration of Primary and Supportive R&D.

A second ongoing effort involves the development of a minimal competency methodology for establishing cutoff scores for the WTPT. Presently, this methodology is being applied to the Jet Engine Mechanic specialty, and cutoff scores have been derived. This research should prove helpful in two main areas. First, information about the percentage of job incumbents tested who are considered minimally competent in a specialty may be useful data for assisting the enlistment standard-setting process. In addition, the training community should benefit from this knowledge when evaluating training effectiveness. A brief explanation of each R&D effort underway or planned can be found in Table 2.

Table 2. Summary of R&D Efforts

Title	Completion schedule	Brief description/major variables
Job Performance Meta-Analysis	FY85	(1) Supervisor, peer, self ratings of job performance data; (2) Multitrait/multimethod (MTMM) analyses of job performance data
WTPT Administrator	FY85	Contractor, active duty (on-site/off-site); rater/ratee familiarity
Rating Form Accuracy	FY86	Purpose of appraisal; rating form specificity; mode and level of instructional detail; accuracy paradigm; acceptability, motivation, confidence, trust
Structural Equation Modeling	FY86	Confirmatory factor analysis of MTMM JPMS rating data; covariance structure analysis of Jet Engine JPMS data
Training Effectiveness Overview	FY86	Air Force training system; use of JPMS for training evaluation
Rater Training	FY86	Type of training; interaction with rating forms; behavior modeling; videotape accuracy; assesment center test bed
WTPT Minimal Competency	FY86	Cut-score methodology; application to Jet Engine Mechanic
Generalizability Theory	FY86	Application to JPMS development, training, administration
Time to Job Proficiency	FY87	Time-based supervisor rating forms; aptitude/experience performance model
WTPT Accuracy	FY87	Error/accuracy interrelationship; rater individual differences
JPMS Technology to Training	FY92	Transition of JPMS methodology to evaluate training effectiveness
JPMS Technology to Officers	FY92	Transition of JPMS methodology to officer jobs
JPMS Technology to Civilians	FY92	Transition of JPMS methodology to validate the civilian Promotion, Placement, and Referral System

VII. SUMMARY

If Air Force use of job performance measures in validating selection procedures proves feasible, a specialty-specific performance measurement database can be kept current and used routinely to validate these selection tests and procedures. The performance measurement technology and resulting data can then be used by: (a) the manpower, personnel, and training R&D community to evaluate their R&D products, (b) the operational manpower and personnel community to evaluate the impact of personnel policies and procedures, and (c) the operational training community to develop and evaluate training.

The Air Force's most promising application of this technology outside of the enlisted selection system project is in training evaluation. The training interface has been specifically requested by Headquarters Air Training Command, and initial investigation of the applicability of this technology to training is underway. The results of the initial research will serve as the guide for further R&D in the training area. If funding permits, plans call for continuation of the training applications through FY92.

REFERENCES

Borman, W.C. (1979). Format and training effects on rating accuracy and rating errors. Journal of Applied Psychology, 64, 410-421.

Cascio, W.F., & Bernardin, H.J. (1981, February). Court cases relevant to employment decisions: Annotated bibliography (AFHRL-TR-80-44, AD-B055 755). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Hedge, J.W. (1984, August). The methodology of Walk-Through Performance Testing. Paper presented at the annual meeting of the American Psychological Association, Toronto.

Kavanagh, M.J., Borman, W.C., Hedge, J.W., & Gould, R.B. (1982). A classification scheme, literature review and directions for research (Contract No. F41689-81-C-0022). Brooks AFB, TX: Manpower and Personnel Division, unpublished manuscript.

Sorcher, M., & Goldstein, A.P. (1972). A behavior modeling approach in training. Personnel Administrator, 35, 35-41.